Contract F61775-98-WE079

SPC 98-4050 –"Simulation of low light level adaptive optics systems"

Final report – August 1999

Dr C.J. Solomon
University of Kent
Canterbury, England

ABSTRACT

Adaptive optics systems are designed primarily for astronomical and surveillance applications to compensate for the effect of random, time-varying aberrations introduced by the earth's atmosphere. Compensation for these aberrations predicates some means of sensing the atmospheric distortion with sufficient signal-noise ratio to apply a meaningful correction. In most scenarios of interest (e.g. remote stellar objects, earth-orbiting satellites), the light level available for this process is low, placing limits on the degree of correction that is possible. The use of artificial guide stars (laser beacons) is one approach to improving the signal-noise available for wavefront sensing. However, the technical effort and cost associated with this solution is very high and still technical problems remain to be solved (e.g. the tilt problem). Certain surveillance applications may also disqualify solutions which employ target illumination. There is thus motivation for the use of systems using simpler technology but employing optimally configured sensing to maximise performance.

The overall study presented is concerned with the development of computer simulation tools enabling the study of optimum imaging scenarios for low light level adaptive optics and of assessing their implications for optical interferometers in which the constituent telescopes employ adaptive optics. The first part of the report outlines the overall scope and scientific basis of an adaptive optics simulation program developed during the study. This includes a description of a graphical-user-interface (GUI) which has been developed to assist the easier use of the program and some explanatory material of the current version of the simulation program. The second part of the report details results of some low light level studies on a 1m telescope employing adaptive optics without laser-guide star technology. The implications for low-light level AO applications are briefly discussed. Part 3 is given in the form of a journal article written in the context of this study and submitted to J.O.S.A. letters entitled – "Variational solution for minimum error norm wavefront projectors". The paper describes the derivation of an optimal method for estimating atmospherically distorted wavefronts by direct integration of the wavefront slope measurements which is also guaranteed to have minimum noise propagation.

1. PHYSICAL BASIS FOR ADAPTIVE OPTICS COMPUTER MODEL

1.1 Software modules

The adaptive optics computer model uses 6 basic software modules which, interfaced in a straightforward manner, simulate the behaviour of an adaptive optics system. The modules correspond to –

- The generation of a sequence of time-evolving atmospheric wavefronts.
- Image formation by a Shack-Hartmann sensor coupled to a CCD camera and estimation of atmospheric wavefront slopes by sub-aperture centroiding.
- Modal estimation of the atmospheric wavefronts.
- Modelling of the spatio-temporal behaviour of the adaptive mirror and thus the drive voltage transfer matrix.
- Implementation of a closed-loop control algorithm, capture of data in the image plane of the telscope and generation of appropriate performance indicators such as time-averaged mean-square error or Strehl ratio.
- A graphical user-interface to allow easy use of the program.

The scientific basis of the computer model is generally accurate and realistic but certain approximations have been made. Accordingly, we first outline the basic principles underpinning our computational implementation of these modules.

1.2 Generation of time-evolving atmospheric wavefronts

The time-evolving wavefront generator is based on the cumulative effect of 4 atmospheric layers, each modelled by a random phase screen. The phase power spectrum is modelled by the Von-Karman spectrum.

The basic approach to simulation is based on random sampling for the phase as a function of space and time such that it will statistically obey the required space-time phase correlation function [1], defined as -

$$\Gamma_{\varphi} = \left\langle \varphi(\vec{x}_i, t_l)\varphi(\vec{x}_j, t_m) \right\rangle \qquad [1.1]$$

Where $\varphi(\vec{x}_i, t_k)$ represents the phase evaluated at point $\vec{x}_i$ and time t_k.
We represent the simulated wavefront by an array of spatial phase values evaluated at N co-ordinates $[\vec{x}_i]_{i=1,N}$ over the pupil. Further, let us assume that we want to generate M such successive phase screens. This sequence of phase screens can be represented by a single vector a of length NxM:

$$a = \left[\varphi(\vec{x}_j, t_l)\right]_{i=1,N*M} \qquad [1.2]$$

We can then write the space time covariance matrix of the phase screens as the covariance of that vector (a matrix of size $(NxM)^2$):

$$\Gamma_{\varphi} = \Gamma_a = \left\langle a_i, a_j \right\rangle \qquad [1.3]$$

The correlation function of the phase is related to the structure function of the phase and can be evaluated for a given model of the refractive index power spectrum such as the Von Karman or Kolmogorov. However, the Kolmogorov spectrum presents a singularity at the spatial frequency zero. It can be shown that the covariance matrix, calculated from this spectrum, will also contain such a singularity. If we considered a piston removed phase, i.e. a phase function with a mean of zero, the covariance matrix does not contain any singularity but still results in an extremely complex expression for the covariance matrix [1]. This singularity does not occur if we assume the Von-Karman spectrum. The covariance matrix corresponding to the Von-Karman spectrum is real, symmetric and positive definite. A Cholesky factorisation of that matrix into two square matrices is therefore possible:

$$\Gamma_a = RR^T \qquad [1.4]$$

The method proceeds by generating a vector of length NxM of independent gaussian random variables b and a product vector c = Rb from which it follows that -

$$\begin{aligned} \mathrm{cov}(c) &= \langle cc^T \rangle \\ &= \langle Rbb^T R^T \rangle \\ &= R\langle \mathrm{bb}^{\mathrm{T}} \rangle R^T \qquad \text{(As R is not a random quantity)} \end{aligned}$$

[1.5]

The components of the Gaussian vector b are statistically independent which means that there is no correlation between them. Further, we may choose them to have unit variance so that the covariance reduces to-

$$\mathrm{cov}(c) = RR^T = \Gamma_a \qquad [1.6]$$

as required.

We have shown above that, by random generation of a vector of unit variance gaussian variables, we can obtain a set of random, temporally correlated phase screens obeying the correct spatio-temporal statistics. This method, however, does require significant computational resources. Consider, for example, if we wish to generate five successive, correlated phase screens of size 128x128 pixels. This implies the need to compute a covariance matrix of size: $(128\text{x}128\text{x}5)^2 = 6710886400 \approx 6.7 * 10^9$ elements. Cholesky factorisation of such a matrix, even by today's standards, requires a very powerful computer. To partly circumvent this problem, we first simultaneously generated only three consecutive screens. Then, by varying the time step between each screen in this set, we may subsequently reconstruct the required sequence of time–evolving phase screens by a temporal interpolation procedure. For example, if we want a sequence of phase screens over a time interval of length T=2s, with a time separation between each screen of $\tau = 2ms$ we first generate a stack of three wavefronts separated by a time step $\frac{T}{2} - \tau$. Then we successively decrease the time step until we produce all the necessary screens. To further conserve computational resources, we may decrease the spatial size of each screen, evaluating the phase values at only a small number of selected spatial co-ordinates. Provided we

use a sampling interval smaller than $\frac{r_0}{2}$, then we can spatially extrapolate the randomly generated phase screens between the sampled points. Details of the technique can be found in reference [1].

1.3 Shack-Hartmann image formation and centroiding

The Shack-Hartmann sensor has been widely studied (see reference [2] for an exhaustive list of references) and is the most common form of wavefront sensor. This is largely due to it's conceptual simplicity and competitive performance. In essence, a Hartmann sensor provides a number of wavefront slope estimates, each averaged over a sub-region of the pupil, by calculation of the barycentre of each sub-aperture image. Our implementation of Shack-Hartmann image formation is a high-fidelity model and includes realistic effects of low light level imaging including both detector and photon noise. The details of the method are outlined in Ash et. al. [3]. The centroiding method employed in our model is based on a recursive centre-of-gravity technique. It has been shown that matched filtering techniques [4] may offer improved slope estimates at low light levels but the current version of the simulation program does incorporate such an estimator.

1.4 Modal estimation of atmospheric wavefronts

This subject has been widely studied. We briefly summarise our approach. The basic method is to expand the random wavefront in terms of a set of chosen basis functions $P_k(\vec{x})$ –

$$\varphi(\vec{x}) = \sum a_k P_k(\vec{x}) \qquad [1.7]$$

Since the Hartmann sensor provides a finite set of local phase gradient measurements, we take the gradient of this equation and evaluate at the coordinates $(\vec{x}_i)$ at which the gradient is known –

$$\nabla\varphi(\vec{x}_i) = \sum a_k \nabla P_k(\vec{x}_i) \qquad (\mathrm{i} = 1,2...\mathrm{M}) \qquad [1.8]$$

This system may be written in matrix form as –

$$\mathbf{Ax} + \mathbf{e} = \mathbf{m} \qquad [1.9]$$

We seek a linear estimator for the unknown modal coefficients given by a linear combination of the slope data -

$$\hat{\mathbf{x}} = \mathbf{Lm} = \mathbf{L}(\mathbf{Ax} + \mathbf{e}) \qquad [1.10]$$

which will minimise the error covariance matrix given by –

$$\mathbf{E} = \left\langle (\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^{\mathrm{T}} \right\rangle \qquad [1.11]$$

Where the matrix L is to be determined. Inserting [1.10] into [1.11] and using the definition of $\mathbf{C}_{\mathbf{x}} = \left\langle \mathbf{x}\mathbf{x}^{\mathrm{T}} \right\rangle$, we obtain -

$$\mathbf{E} = [\mathbf{I} - \mathbf{LA}]\, \mathbf{C}_{\mathbf{x}} [\mathbf{I} - \mathbf{LA}]^{\mathrm{T}} + \mathbf{LVL}^{\mathrm{T}} \qquad [1.12]$$

Taking first variations in $\mathbf{E}$ with respect to $\mathbf{L}$ and setting to zero yields –

$$\delta\mathbf{E} = \delta\mathbf{L}(\mathbf{AC}_{\mathbf{x}}[\mathbf{I} - \mathbf{LA}]^{\mathrm{T}} + \mathbf{VL}^{\mathrm{T}}) + ([\mathbf{I} - \mathbf{LA}]\mathbf{C}_{\mathbf{x}}\mathbf{A}^{\mathrm{T}} + \mathbf{LV})\delta\mathbf{L}^{\mathrm{T}} = \mathbf{0} \qquad [1.13]$$

The optimal matrix is thus given by -

$$\mathbf{L}_{MAP} = \mathbf{C}_x\mathbf{A}^T\left(\mathbf{A}\mathbf{C}_x\mathbf{A}^T + \mathbf{V}\right)^{-1} \qquad [1.14]$$

This may also be written as -

$$\mathbf{L}_{OPT} = \left(\mathbf{A}^T\mathbf{V}^{-1}\mathbf{A} + \mathbf{C}_x^{-1}\right)^{-1}\mathbf{A}^T\mathbf{V}^{-1} \qquad [1.15]$$

Substituting this optimal (MAP) estimator into the definition of the error covariance matrix (eq. 1.10) gives the expression -

$$\mathbf{E}_{OPT} = \left(\mathbf{A}^T\mathbf{V}^{-1}\mathbf{A} + \mathbf{C}_x^{-1}\right)^{-1} \qquad [1.16]$$

• If we assume that we have no prior knowledge of the modal coefficients (i.e. $\mathbf{C}_x \to \infty$), this optimal estimator reduces to the well-known BLUE estimator –

$$\mathbf{L}_{OPT} = \left(\mathbf{A}^T\mathbf{V}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{V}^{-1} \qquad [1.17]$$

If further we may assume that the errors on the slope measurements have equal variance and are uncorrelated (i.e. $\mathbf{V} = \sigma^2\mathbf{I}$) we obtain a standard least-squares solution –

$$\mathbf{L}_{OPT} = [\mathbf{A}^T\mathbf{A}]^{-1}\mathbf{A}^T \qquad [1.18]$$

We have implemented modal wavefront estimation in our model using both the standard least-squares solution (eq. [1.17]) and the optimal MAP estimator (eq. [1.18]). The current version of the software uses the well-known Zernike circular polynomials [5] as a basis set.

1.5 Adaptive mirror

The simulated wavefront correction device was based on an independent tip-tilt mirror and a 17 element bimorph mirror. The influence functions used to represent the action of the mirror electrodes follow the theoretical model of Halevi [6]. This model assumes that the mirror is composed of one piezoelectric plate glued with a plate of mirror glass activated by a distribution of electrodes placed on the reverse side of the mirror. The device has a rectangular shape of size ab m^2. In such a model, the vertical displacement of the mirror W(x,y) and the voltage distribution V(x,y) must satisfy the following differential equation -

$$\nabla^4 W + A\nabla V^2 = 0 \qquad [1.19]$$

The coefficient A is a constant proportional to $d_{13}t_m^{-2}$, where d_{13} is a coefficient of the piezoelectric tensor[i] and t_m is the thickness of the mirror. Typically A is of magnitude 10^{-9} $m^{-1}V^{-1}$. The model assumes that the mirror is constrained at the edges (*i.e.* W=0 at the edges). Then, by representing the voltage distribution as a double Fourier series, and inserting it into eq. [1.19], it follows that the vertical displacement can be expressed [6] as -

$$W(x,y) = \sum_{n=1}^{+\infty}\sum_{m=1}^{+\infty}\frac{AV_{nm}}{\pi^2\left(n^2/a^2 + m^2/b^2\right)\sin\left(\frac{n\pi x}{a}\right)\sin\left(\frac{m\pi y}{b}\right)}$$

$$V_{vnm} = \frac{4}{ab}\int_0^a\int_0^b V(x,y)\sin\left(\frac{n\pi x}{a}\right)\sin\left(\frac{m\pi y}{b}\right)dxdy$$

As this model of the mirror was square, we use a circular mask to match the mirror to the conjugate image of the pupil (therefore constraining it at the edge of the mask). The distribution of the electrodes was chosen as a double circular ring. The model assumes that the mirror is linear in it's response and that it exhibits no hysteresis, assumptions which are not exactly fulfilled in practice but which nonetheless do not generally exhibit a major influence on performance. The time response of the mirror is not modelled but it's overall effect on imaging can effectively be included by introducing an overall time-delay between calculation of the required drive voltages and their application to the actuators and this is included in the simulation.

1.6 Closed-loop control algorithm

Our computer model allows the use of either an *integral control* algorithm or a *Smith compensator* [7]. In both cases, the control system applies a correction to the bimorph mirror device every τ seconds and the mirror profile is maintained during this interval τ until the next correction is made. The correction to the mirror comprises an adjustment to the voltages applied to the actuators and is calculated on the basis of the difference between the residual wavefront as measured by the wavefront sensor and the profile currently existing on the adaptive mirror.

Let the set of voltages $y_k(n(\tau-1))$ be the voltages existing on the actuators at time $n(\tau-1)$. The voltages applied to the actuators at time $n\tau$, $y_k(n\tau)$ are given by the integral recursion –

$$y(n\tau) = y((n-1)\tau) + K_c\,\tau\,u(n\tau) \qquad [1.21]$$

where the set of modification voltages u(nτ) are calculated from the present slope data provided by the wavefront sensor and K_c is the gain of the integral controller. The time delay τ and the gain of the integral controller may be adjusted by the user.

The Smith controller was proposed to compensate for the problem of time delay in automatic control systems in 1956 by O.J. Smith. At that time, digital electronics was not sufficiently developed to allow it's easy implementation. It is now, however, a well-used method in control systems as a device primarily added to the integral controller to increase stability. The approach consists in adding a feed-back loop, which simulates the behaviour of the plant but without the correction time delays resulting from finite calculation and response time during the data processing and control actuation. If we let the set of voltages y(n(T-1)) be the voltages existing on the actuators at time n(T-1), for the Smith compensator the voltages applied to the actuators at time nT, y(nT) are given by the integral recursion –

$$y(nT) = \frac{KT}{1-KT}u(nT) + \frac{y((n-1)T)}{1-KT} - \frac{KT}{1-KT}y((n-2)T) \qquad [1.22]$$

where the set of modification voltages u(nT) are calculated from the present slope data provided by the wavefront sensor and K is the gain of the controller. The time delay T and the gain of the integral controller may be adjusted by the user.

The Smith controller reduces then to a second order algorithm as opposed to the first order algorithm of the integral controller.

1.7 Graphical User Interface (GUI)

The aim of the GUI is to effect an easy and convenient interface for the user to carry out simulations of adaptive optics imaging scenarios. A related aim was to construct the interface in such a way as to reduce the possibility of unintentional input/errors. A GUI has been developed which will allow the input of key simulation parameters, allow the easy visualisation and storage of results. The GUI is effective in its basic aims though there is scope for more flexibility.

1.8 Current restrictions on the adaptive optics simulation

In it's current form, there are certain restrictions on the type of adaptive optics system and the range of operating conditions that can be modelled.

- The program considers only a Shack-Hartmann sensor with 52 active sub-apertures. Note that the current version does, however, allow variation in the sub-aperture size expressed in terms of the Fried parameter r_0. Critically, this does allow the subaperture size to r_0 ratio to be varied, a parameter which is known to significantly affect low light level performance [3,8].
- The basis matrix for estimation of atmospheric wavefronts is modelled in terms of the Zernike polynomials only. This is not a serious restriction but a future version should allow the use of the Karhunen-Loeve basis.
- The adaptive mirror is a bimorph device with 17 actuators. There is no facility to consider other types of correction device.
- Although a variety of closed-loop control algorithms are possible, in principle, the current version permits an integral controller or a Smith compensator only.
- Basic program output is the time-averaged, compensated Strehl ratio in the image plane as a function of the controller gain. This is a satisfactory measure of the basic performance. Program output for intermediate stages such as random phase screens and Hartmann data and image plane data is possible but can not currently be effected from the GUI.

1.9 Compiling and executing the simulation code.

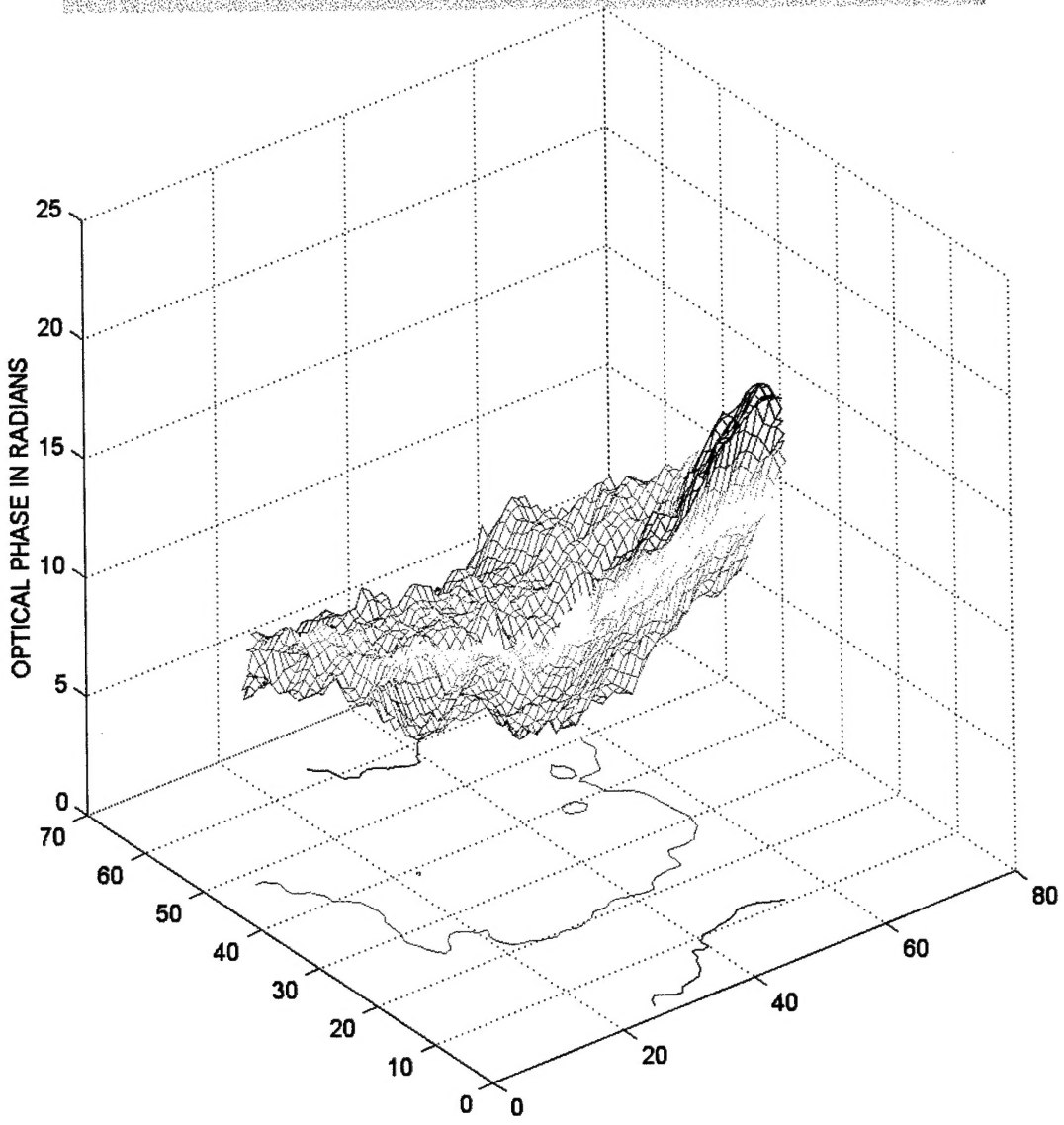
The simulation source code consists of nearly 10 000 lines of C. A number of data structures are also necessary for the program to run. If at least one pre-calculated set of time-evolving wavefronts are included this occupies approximately 100 Mbytes.

- The source code files (which are too numerous to list) can be transferred on request to users as required by ftp.
- All source code files should be placed in a single directory.
- To compile the program, an ANSI C compiler is required. A text file *prbc.txt* contains the necessary commands to compile and link the code.

1.10 Using the program

Within the constraints specified in section 1.8, the user may select a number of observational and system parameters which may be expected to influence the overall performance of the AO system using a GUI. The start-up GUI is shown overleaf -

LOW FLUX AO SIMULATOR



Enter flux (photoelectrons /m^2)	2500
Enter Fried's parameter in cm	20
Enter telescope diameter (in metres)	1.2
Enter correction time delay in ms	2.5
Enter CCD read-noise rms PEs/pixel/frame	3.0
SELECT CONTROLLER	Integral / Smith
ENTER GAIN	200

PROCESS SAVE DATA

DISPLAY DATA EXIT

COMMAND LINE INTERFACE

The program may also be run using a simple command line menu which asks the user to specify the conditions of the simulation. A typical example is given in figure 1 and an explanation of the choices made by the user follows.

>>Do you want to use pre-set values for the telescope diameter and turbulence conditions (i.e. D, r0, v..) ? (y/n)

This is an important choice because the calculation of a sequence of statistically correct, time-evolving atmospheric wavefronts takes considerable computational resources and a long time to generate (many hours). If you have a pre-calculated set of wavefronts corresponding to a given telescope diameter D and Fried parameter r_0 which you wish to use, you should answer **n.** In the example in figure 1, the user has answered **n.**

>>Enter filename containing this information.

Enter the filename containing the telescope diameter and turbulence parameters for which the time-evolving wavefront is already pre-calculated.

>>Do you want to use a set of system parameters already existing ? (y/n)

Here, the user may decide whether to simulate an imaging scenario for which the relevant parameters are prespecified in a file or alternatively specify new system parameters which will be written to file themselves. In the example in figure 1, the user has answered **n.**

>>Enter file name for saving new simulation parameters

An appropriate file-name is entered.

>>Infinite or finite light level? (1 for ∞: 2 for finite)

The object being imaged must be specified as being very bright (i.e. providing an essentially ∞ light level at the wavefront sensor) or a finite light level. In the example in figure 2, the user has specified a finite light level.

>>Enter visual magnitude of the natural guide star.

Give the visual magnitude of the natural guide star being used for wavefront sensing

>>Enter sensitive bandwidth of wavefront sensor camera $\Delta\lambda$ (nm)

The user must specify the optical wavelength region in nanometres to which the wavefront sensor camera is sensitive.

>>Enter CCD quantum efficiency (0 and 1)

Specify the quantum efficiency of the CCD camera

>>Specify level of CCD noise (rms photoelectrons/per pixel/frame)

The user must specify the noise performance of the CCD camera in terms of the rms photoelectrons generated per pixel per camera frame.

>>Enter minimum gain for integral controller: kmin

The output of the AO simulation is the time-averaged, compensated Strehl ratio as a function of the gain of the integral controller. Specify the minimum gain of interest.

>>Enter maximum gain for integral controller: kmax

The output of the AO simulation is the time-averaged, compensated Strehl ratio as a function of the gain of the integral controller. Specify the maximum gain of interest.

>>Enter overall time delay between wavefront measurement and correction on mirror (ms)

Any real adaptive optics system operating in closed-loop will introduce a delay between calculating the current form of the residual wavefront and applying an appropriate correction to the mirror. Specify this delay in ms.

>>Enter numeric label for storing result of simulation

This allows the user to label the simulation results file via a parameter *c_offset* which is set

>>Enter any 5 digit random number (for noise simulation)

This ensures that each simulation will begin with a different random number seed.

>>Enter directory label

This allows the user to specify the directory in which to place the simulation results file. The parameter *dir_hold*

The simulation is carried out and the results are stored in a file named ~/*/zern_result*.d. The first asterisk corresponds to the directory label *dir_hold* and the second to the numeric label *c_offset*. In other words, if the program is executed from the directory "current_working_directory", choosing *dir_hold = myplace and c_offset = 3* will cause the output file to be written as - ~/current_working_directory/myplace/zern_result3.d .This file contains the compensated Strehl ratios achieved for each value of the integral gain

2 Low-light level simulations – some preliminary results

In this section, we present results of the performance of a Shack-Hartmann based adaptive optics system operating under low photon flux conditions. The new generation of large, ground-based telescopes in the 5-10 m diameter range which are currently being built and commissioned in many parts of the world (e.g. ALFA, GEMINI, SUBARU) will all be equipped with adaptive optics systems employing laser-guide stars. We thus stress that the relevance of the presented results lies more in the use of smaller ground-based telescopes (possibly several small telescopes employed in an interferometric configuration) where more modest D/r0 values prevail and laser-guide star technology is either not available or undesirable.

The series of simulations assumed the following **fixed conditions** –

1. D/r0 = 5. Telescope diameter 1m and r0 = 20cm.
2. Average wind speed of 20 m/s. For a Fried parameter of r0=20cm, this gives a Greenwood time-delay of 3.15 ms.
3. A Shack-Hartmann wavefront sensor having 52 active sub-apertures.
4. A CCD camera of fixed quantum efficiency.
5. A 17 actuator bimorph mirror.

The following **conditions were varied** –

The incident photoelectron flux.
The read-noise in photoelectrons per pixel per frame on the CCD camera.
The time-lag τ between wavefront sensing/estimation and application of the actuator signals
The gain on the closed-loop control algorithm (either the integral controller or the Smith Compensator) was varied.

The aim was to study the mean compensated Strehl ratio of the simulated AO system as these parameters varied. In particular, the optimisation of the closed-loop control gain was studied for several different signal-noise/correction time-delay scenarios.

Figure 1 shows the mean-compensated Strehl ratio as a function of controller gain. In this study, the incident flux was high (i.e. assumed to be infinite). The control algorithm was the integral controller and the correction time-lag τ was set = 1.5 ms. This time-lag is only approximately 1/3 of the Greenwood time delay and the system is thus fast enough to accurately track the atmospheric fluctuations. This rather idealised case allowed us to test the limiting behaviour of the integral controller. We note an optimum gain around 400 (this corresponds to a linearisation factor $K_I\tau \sim 1$). As we go to higher values of gain the system becomes unstable due to the spatial modelling error and the fact that the time delay, although small, is still finite.

Figure 2 shows the mean-compensated Strehl ratio as a function of controller gain. In this study, the incident flux was low, corresponding to a mean value of 60 photoelectrons/sub-aperture. The control algorithm was again the integral controller and the correction time-lag τ = 1.5 ms. A read-noise level of 1.5 photoelectrons per pixel per frame (state-of-the-art devices currently achieve

values of ~ 3) was also included. The main features of figure 2 are largely in agreement with intuition. Overall compensated Strehl is reduced and the gain curve is more peaked with a maximum occurring around $K_I \approx 350$. The rapid decrease in performance above the optimum reflects the tendency of the integral controller to amplify the noise.

Figure 3 was obtained under identical conditions to figure 2 except that the Smith controller was implemented. In this case, we see that the optimum gain is somewhat lower and overall performance slightly poorer for high gain. As the correction time delay is much smaller than the Greenwood time, we do not expect any significant improvement from the elimination of the time delay and the poor noise propagation for high gain results in a slightly lower performance.

Figures 4 and 5 show the mean-compensated Strehl ratio as a function of controller gain for integral and Smith controllers with a large correction time delay of 8ms (> 2.5 times the Greenwood time). The flux was 100 photoelectrons per sub-aperture and the read-noise at a level of 1.5 photoelectrons/sub-aperture/frame. Despite the increase in signal-noise, figures 4 and 5 show a decrease in the overall compensation as compared to figures 2 and 3. This results from the significant correction time delay. We note that in this case, the Smith compensator performs slightly better than the integral controller and is stable across a larger range of gains.

These studies have demonstrated the ability of the program to simulate AO systems operating in low light level regimes. Significant improvements in Strehl ratio over the uncompensated state (up to S= 0.13) can be achieved even at flux levels of 60 photoelectrons/sub-aperture using AO compensation which uses a very low-noise CCD camera (1.5 electrons rms/pixel/frame) in the wavefront sensing arm. We have also demonstrated the ability of the Smith controller to partially compensate the correction time-delay and yield improved performance over the standard integral controller. A critical aspect in these low flux studies was the read-noise performance of the simulated CCD camera – a factor of 3 increase in the read-noise under the same conditions was sufficient to effectively reduce AO compensation to a performance level similar to uncompensated imaging. So, while we stress that greater sensitivity in the wavefront sensing can lead to enhanced performance and can be achieved via several strategies that have been studied in related work (namely, matched filtering techniques for centroiding [4], optimal Bayesian wavefront estimators [9] and also the crucial choice of the Hartmann sub-aperture size to match the prevailing signal-noise level [3], the importance of using low noise imaging devices in the wavefront sensing arm must be emphasised. As current state-of-the art CCD devices yield read-noises of the order of 3 electrons rms/pixel/frame, photon counting devices which have effectively no noise (but currently considerably lower quantum efficiency) may prove to be the right instrument of choice for low light level AO. A simulated photon counting device having 20 % quantum efficiency but no read-noise yields a broadly similar performance to that given in figure 1. A systematic comparison between such devices, however, remains to be carried out.

FIGURE 1: MEAN COMPENSATED STREHL RATIO / GAIN

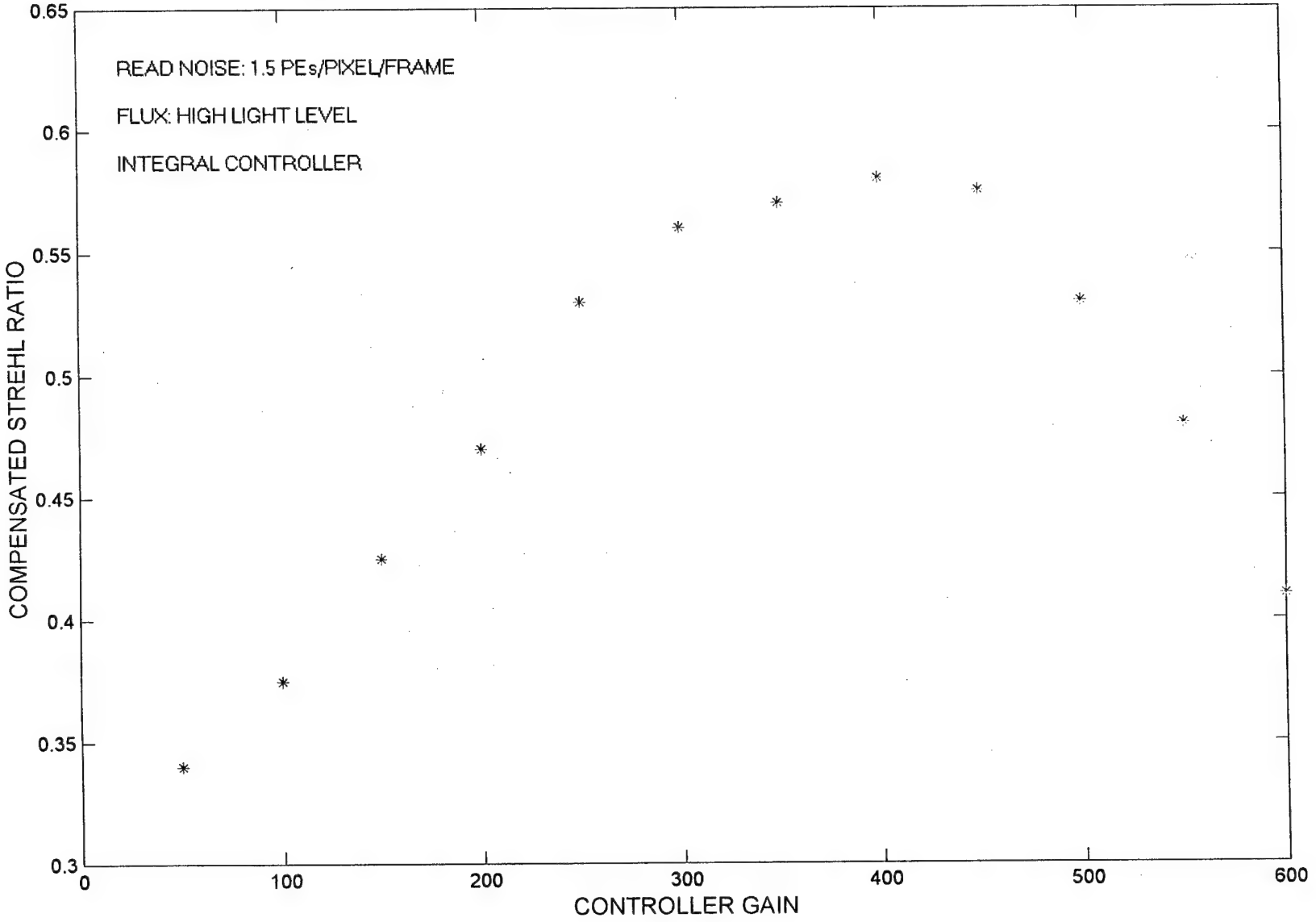


FIGURE 2: MEAN COMPENSATED STREHL RATIO / GAIN

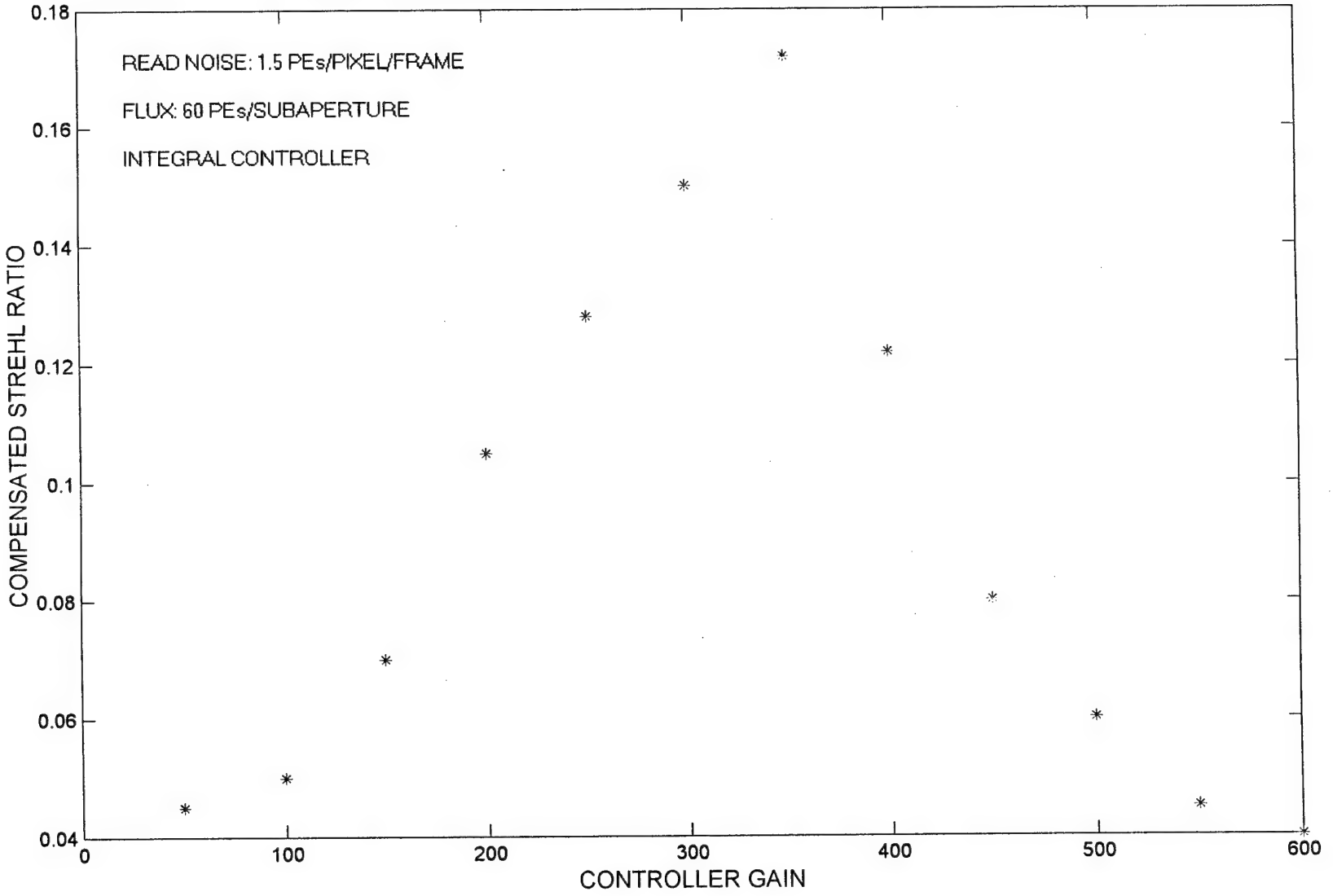


FIGURE 3: MEAN COMPENSATED STREHL RATIO / GAIN

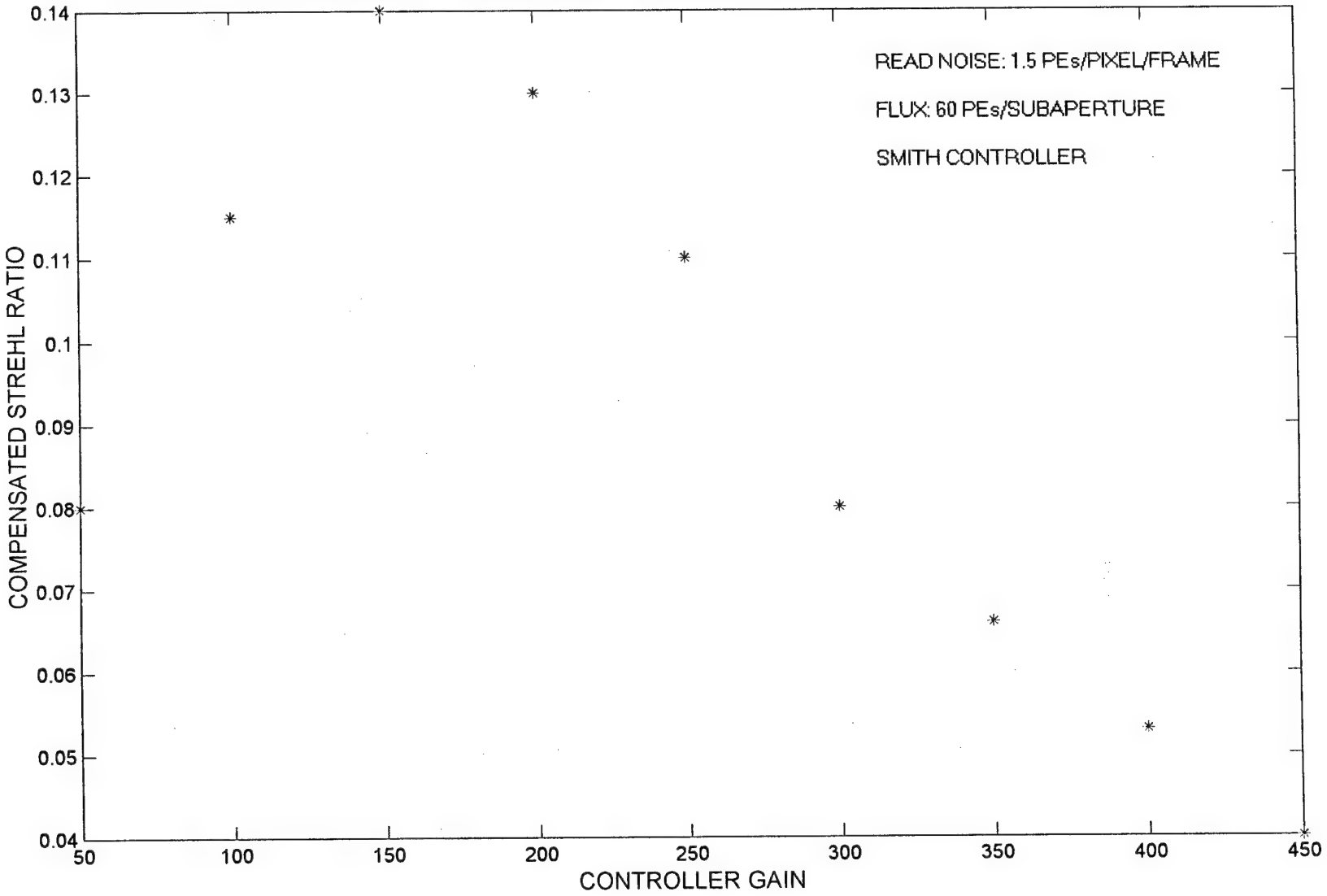


FIGURE 4: MEAN COMPENSATED STREHL RATIO / GAIN

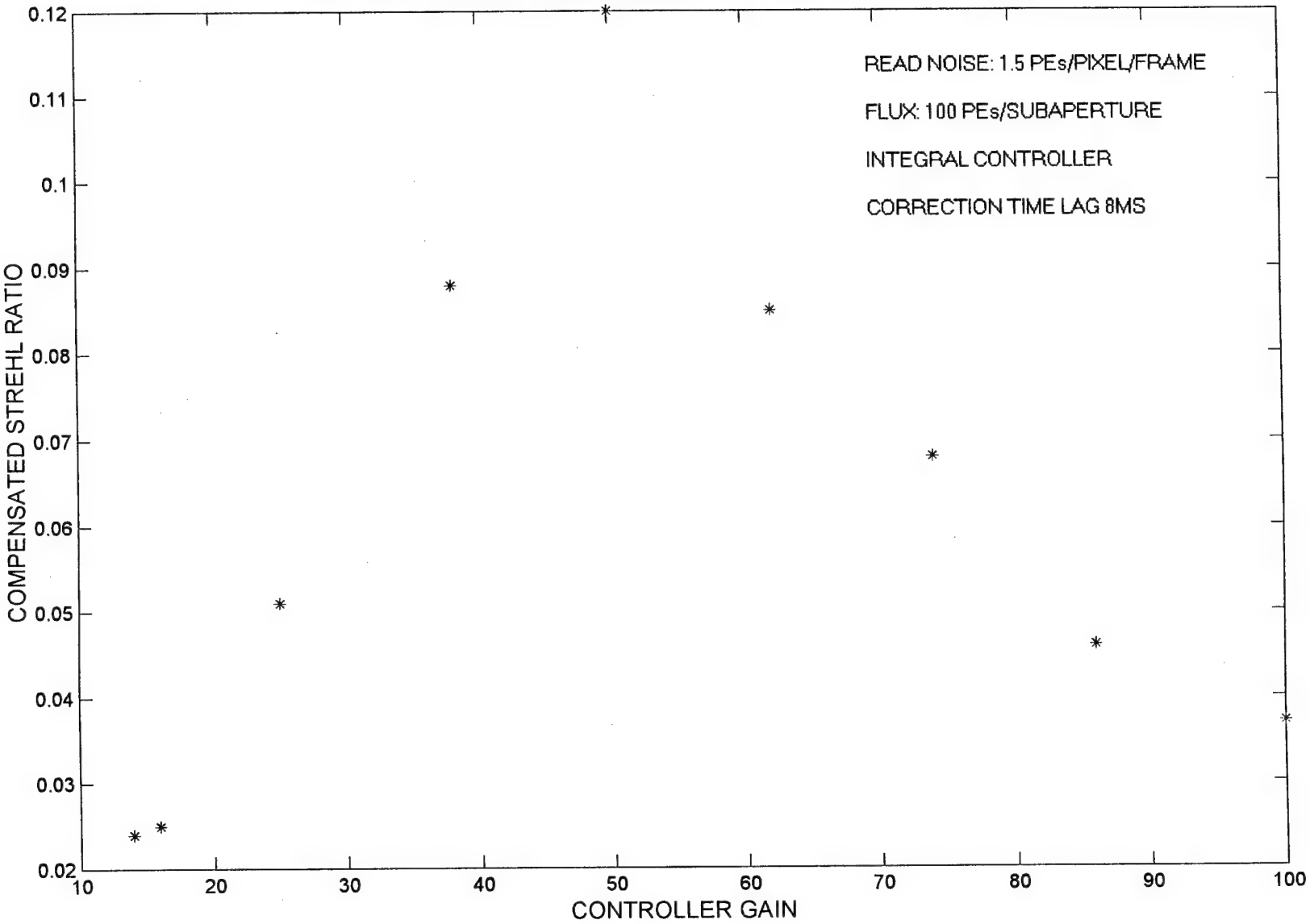
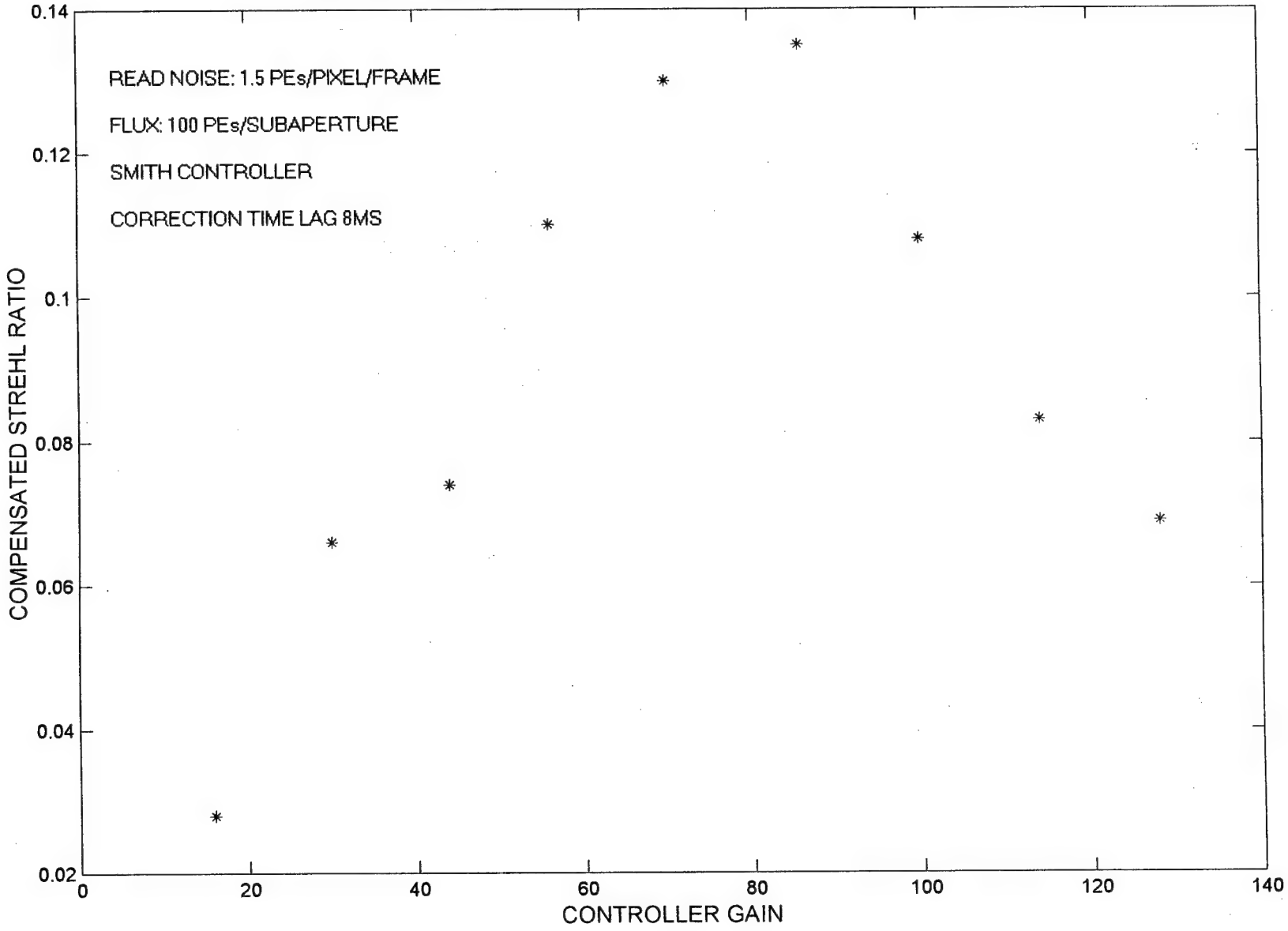


FIGURE 5: MEAN COMPENSATED STREHL RATIO / GAIN



Section 2: Summary

- The software is able to simulate the imaging capacity of an AO system viewing an object confined to a region within a single isoplanatic patch. The predominant interest is in using this system for low light level studies. No laser guide star system is assumed and no science-object is necessary to assess performance. The AO system consists of a 17 element bimorph adaptive mirror coupled to a Hartmann sensor employing 52 active sub-apertures and a low-noise CCD camera operating in closed-loop with an integral or Smith controller. The *incident flux from the wavefront sensor object, turbulence conditions, telescope diameter, type and gain of the controller, correction time delay and CCD camera noise performance* may all be adjusted by the user. Preliminary results suggest that the software performs correctly and have enabled basic low light level imaging performance to be assessed as a function of the variable parameters.

Future improvements could be made in two main areas–

- Incorporating greater flexibility into the GUI to control a larger range of study/observational condition.
- Addition of software modules to include different adaptive mirrors or correction devices (e.g. LCs), different imaging devices such as photon-counting cameras and more sophisticated approaches to wavefront estimation and control (e.g. Matched filtering techniques for centroid estimation and predictive control algorithms).

References

1. Roggemann M.C., "Limited degree of Freedom adaptive optics and image reconstruction", Appl. Opt. 30 (29) pp-4227-4233 (1991).
2. Tyson R., "Principles of Adaptive Optics", S.P.I.E Press (1991).
3. Ash D.L., Solomon C.J. and Loos G. , "Optimal Hartmann sensing at low light levels", Opt. Comm. 156 pp10-15 (1998).
4. Ruggiu J.M., Solomon C.J. and Loos G., "Gram-Charlier matched filter for Shack-Hartmann sensing at low light levels", Opt. Lett. 23, no. 4, pp235-237 (1998).
5. Noll R.J. "Zernike polynomials and atmospheric turbulence", J. Opt. Soc. Am. A, **66** (207) (1976)
6. Halevi P., J. Opt. Soc. Am. A, **73**, pp110-116 (1982)
7. Franklin G.F., Powell J. D. and Workman M.L, "Digital control of dynamic systems", 2nd edition, Addision-Wesley (1994).
8. Gavel D.T., Morris J.R. and Vernon R.G., "Systematic design and analysis of laser-guide star adaptive optics systems for large telescopes", J. Opt. Soc. Am. A, **11** (2)(1994)
9. "C.J. Solomon, N.J. Wooder and J.C. Dainty, "Bayesian estimation of atmospherically distorted wavefronts using Shack-Hartmann sensors" (Opt. Rev. Vol. 2, No. 3, 1995)

[i] E. Steinhaus & S.G. Lipson, *J. Opt. Am. A.*, **69**, p. 478 (1979)

VARIATIONAL SOLUTION FOR MODAL WAVEFRONT PROJECTORS OF MINIMUM ERROR NORM

Christopher J. Solomon

† School of Physical Sciences, University of Kent, Canterbury CT2 7NR, U.K.

Gary C. Loos

Air-Force Research Laboratories, Kirtland Air-Force base, Albuquerque, NM 87117-6008, U.S.A.

Susana Rios

Area de Optica, Univ. Santiago de Compostela, E-15706 Compostela, Spain

Abstract

Common wavefront sensors such as the Hartmann or curvature sensor provide measurements of the local gradient or Laplacian of the wavefront. Expression of wavefronts in terms of a set of orthogonal basis functions thus generally leads to a linear wavefront estimation in which modal cross-coupling occurs. Auxiliary vector functions may be derived which effectively restore the orthogonality of the problem and enable the modes of a wavefront to be independently and directly projected from slope measurements. Using variational methods, we derive the necessary and sufficient condition for these auxiliary vector functions to have minimum error norm. For the specific case of a slope-based sensor and a basis set comprising the Zernike circular polynomials, these functions are precisely the Gavrielides functions.

In problems of wavefront estimation, it is common to expand the wavefront in terms of a set of orthogonal functions or modes -

$$\Phi(\underline{x}) = \sum_{k=1}^{N} a_k P_k(\underline{x}) \tag{1.1}$$

which obey the orthogonality relation -

$$\int_D P_i(\vec{x}) P_j(\vec{x}) d^2x = \delta_{ij} \tag{1.2}$$

and D denotes the domain of integration.

The most commonly used set of functions is the Zernike circular polynomials. Speaking generally, orthogonal bases are desirable because they allow the modal coefficients to be evaluated by simple integration of a product of two functions over the domain, using the orthogonality relation eq. 1.2. However, since Hartmann sensors (and shearing interferometers) provide estimates of the *gradient of the phase* rather than the phase itself, the appropriate model is -

$$\langle \nabla \Phi(\underline{x}_j) \rangle = \sum_{k=1}^{N} a_k \langle \nabla P_k(\underline{x}_j) \rangle \tag{1.3}$$

where the coordinates $\underline{x}_j$ denote the position of the slope measurements within the pupil. In this case, the orthogonality of the basis cannot be exploited. Modal cross-coupling will occur and the coefficients must be obtained by solving an inverse problem/overdetermined system of linear equations. Many workers have examined this problem in a search for optimal solutions [1-6].

The evaluation of modes by direct integration could be restored if we can derive a set of auxiliary vector functions $\mathbf{F}_i(\underline{x})$ which are *orthogonal to the gradients of the basis functions*. To establish the necessary conditions for such a set to exist, we consider that these functions obey the relation -

$$\vec{\nabla} \cdot \mathbf{F}_i(\underline{x}) = P_i(\underline{x}) \tag{1.4}$$

Using the orthogonality relation eq. 1.2 in eq. 1.1 and substituting eq. 1.4, the modal coefficients are given by -

$$a_i = \int_D \phi(\underline{x}) P_i(\underline{x}) d^2x = \int_D \phi(\underline{x}) \vec{\nabla} \cdot \mathbf{F}_i(\underline{x}) d^2x$$

This may be written as -

$$a_i = \int_D \vec{\nabla} \cdot (\mathbf{F}_i(\underline{x}) \phi(\underline{x})) d^2x - \int_D \vec{\nabla} \phi(\underline{x}) \cdot \mathbf{F}_i(\underline{x}) d^2x$$

Applying the divergence theorem, we obtain -

$$a_i = \oint_C \phi(\underline{x})\mathbf{F}_i(\underline{x}) \cdot \vec{dl} - \int nabla\phi(\underline{x}) \cdot \mathbf{F}_i(\underline{x}) d^2x \tag{1.5}$$

Clearly if the required set of functions $\mathbf{F}_i(\underline{x})$ satisfies the relations -

$$\vec{\nabla} \cdot \mathbf{F}_i(\underline{x}) = P_i(\underline{x})$$

$$\mathbf{F}_i(\underline{x}) \cdot d\underline{l} = 0 \tag{1.6}$$

where $F_i(\underline{x}) \cdot d\underline{l}$ denotes the normal component of the function to the closed contour C of the integration domain D, we then have -

$$a_i = -\int_D \nabla\phi(\underline{x}) \cdot \mathbf{F}_i(\underline{x}) d^2x \tag{1.7}$$

In other words, we may evaluate the modes by *direct integration* as required. It has been shown that there are, in fact, a number of possible sets of vector polynomials which may be used in eq. 1.7 [7-9] and there is no unique solution to the problem described by eq. 1.6.

In the remainder of this letter, we will address the question of how to derive the optimum set of vector functions $\mathbf{F}_i(\underline{x})$. By optimum, we mean that set of vector functions $\mathbf{F}_i(\underline{x})$ which will give the minimum error or noise propagation in our estimate of the wavefront given by eq. 1.7.

The estimate of the k^{th} modal coeffficient is given by -

$$\widehat{a}_k = -\int_D \mu(\underline{x}) \cdot \mathbf{F}_k(\underline{x}) \, d^2\underline{x} \tag{1.8}$$

where $\mu(\underline{x}) = \overrightarrow{\nabla}\phi(\underline{x}) + \nu(\underline{x})$, and $\nu(\underline{x})$ is the additive noise vector at $\underline{x}$. If we may assume a noise process having zero mean and covariance $\sigma_\nu^2\delta(\underline{x} - \underline{x}')$ (which is reasonable for Hartmann-type sensors), it is straightforward to show that the variance in radians^2 associated with our estimate of the k^{th} mode is

$$N(\mathbf{F}_k) = \sigma_\nu^2 \int_D |\mathbf{F}_k(\underline{x})|^2 \, d^2x \tag{1.9}$$

The noise propagator is thus defined as

$$N(\mathbf{F}_k) = \int_D |F_k(\underline{x})|^2 \, d^2x \tag{1.10}$$

This noise propagator thus depends *only* on the particular choice of the vector functions through the volume defined by $|\mathbf{F}_k^2(\underline{x})|$in D. The ensemble-averaged mean-square error associated with the estimator is then -

$$\sigma^2 = \sigma_\nu^2 \sum_{k=2}^{M} N_p(\mathbf{F_k}) + \sum_{k=M+1}^{\infty} \langle a_k^2 \rangle \tag{1.11}$$

where the latter term simply corresponds to unestimated modes. As the conditions specified in eq. 1.6 do not determine a unique solution for $\mathbf{F}_k(\underline{x})$, there are many possible solutions for each mode, each having a different associated noise propagator.

The key point is that we seek to minimise the noise propagators expressed by eq. 1.10 subject to the constraint equation and boundary condition expressed by eq. 1.6. This is a variational problem. Accordingly, we introduce the Lagrange multiplier function $\lambda_k(\underline{x})$ and seek to minimise an objective function given by -

$$Q = \int \mathbf{F}_k^2(\underline{x}) d^2x + \int \lambda_k(\underline{x})[\nabla \cdot \mathbf{F}_k(\underline{x}) - P_k(\underline{x})] d^2x \tag{1.12}$$

We now take first variations in Q with respect to $\mathbf{F}_k(\underline{x})$ and set these to zero to find the stationary points of Q -

$$\delta Q = 2 \int \delta \mathbf{F}_k(\underline{x}) \cdot \mathbf{F}_k(\underline{x}) d^2x + \int \lambda_k(\underline{x}) \nabla \cdot \delta \mathbf{F}_k(\underline{x}) d^2x \quad = 0 \tag{1.13}$$

Since $\nabla \cdot [\lambda_k(\underline{x}) \mathbf{F}_k(\underline{x})] = \lambda_k(\underline{x}) \nabla \cdot \mathbf{F}_k(\underline{x}) + \mathbf{F}_k(\underline{x}) \cdot \nabla \lambda_k(\underline{x})$ this may be expressed as -

$$\delta Q = 2 \int \delta \mathbf{F}_k(\underline{x}) \cdot \mathbf{F}_k(\underline{x}) d^2x + \int \nabla \cdot [\lambda_k(\underline{x}) \delta \mathbf{F}_k(\underline{x})] d^2x - \int \delta \mathbf{F}_k(\underline{x}) \cdot \nabla \lambda_k(\underline{x}) d^2x \quad = 0 \tag{1.14}$$

Applying the divergence theorem to the second integral in eq. 1.14 and rearranging we obtain -

$$\int [2\mathbf{F}_k(\underline{x}) - \nabla \lambda_k(\underline{x})] \cdot \delta \mathbf{F}_k(\underline{x}) d^2x + \int_C \lambda_k(\underline{x}) \delta \mathbf{F}_k(\underline{x}) \cdot d\underline{l} \quad = 0 \tag{1.15}$$

The boundary condition $\int \mathbf{F}_k(\underline{x}) \cdot d(\underline{l}) = 0$ from eq. 1.6 ensures that the last integral in eq. 1.15 is identically zero and we arrive at the simple condition for the $\mathbf{F}_k(\underline{x})$ to have minimum error norm -

$$\mathbf{F}_k(\underline{x}) = \nabla \lambda_k(\underline{x}) \tag{1.16}$$

In other words, the $\mathbf{F}_k(\underline{x})$ are an *irrotational* set of functions - i.e. must be expressible as the gradient of an associated scalar function. Note we have dropped the factor of 2 as this will have no effect on the functional form for $\nabla \lambda_k(\underline{x})$). Substituting eq. 1.16 into the constraint eq 1.6 we find that the required scalar field $\lambda_k(\underline{x})$ is determined by the Poisson equation with Neumann boundary condition -

$$\nabla^2 \lambda_k(\underline{x}) = P_k(\underline{x}) \quad \underline{x} \in D$$

$$\vec{\nabla}\lambda_k(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0 \quad \mathbf{r} \in C \tag{1.17}$$

This is the main result of our analysis. *The Poisson equation with Neumann boundary conditions has a unique solution and the minimum error norm, modal projector functions can be obtained by solving eq. 1.17 for the given set of basis functions,* $P_k(\underline{x})$. Such a set of functions was, in fact, explicitly derived in analytic form for the Zernike circular polynomial basis by Gavrielides in 1982 [10] although the noise propagation behaviour of these functions was not addressed in his original work. For completeness, their analytic form is given here -

The components of the Gavrielides vector polynomials are defined in polar coordinates as follows -

For $m \neq 0$

$$F_r^i = \frac{1}{\pi}\sqrt{2(n+1)}T_n^m(r)cosm\theta \text{ if i even}$$

$$F_r^i = \frac{1}{\pi}\sqrt{2(n+1)}T_n^m(r)sinm\theta \text{ if i odd}$$

$$F_\theta^i = \frac{1}{\pi}\sqrt{2(n+1)}Q_n^m(r)[-msinm\theta] \text{ if i even}$$

$$F_\theta^i = \frac{1}{\pi}\sqrt{2(n+1)}Q_n^m(r)[mcosm\theta] \text{ if i odd}$$

where

$$T_n^m(r) = \frac{1}{4}\sum_{s=0}^{\frac{n-m}{2}} \frac{C_n^m(s)\,(n-2s+2)\left[r^{m-1} - r^{n-2s+1}\right]}{\left(\frac{n+m}{2}-s+1\right)\left(\frac{n-m}{2}-s+1\right)}$$

$$Q_n^m(r) = \frac{1}{4m}\sum_{s=0}^{\frac{n-m}{2}} \frac{C_n^m(s)\left[(n-2s+2)\,r^{m-1} - mr^{n-2s+1}\right]}{\left(\frac{n+m}{2}-s+1\right)\left(\frac{n-m}{2}-s+1\right)}$$

and where $C_n^m(s) = \frac{-1^s(n-s)!}{s!(\frac{n+m}{2}-s)(\frac{n-m}{2}-s)}$

For m=0

$$F_r^i = \frac{1}{\pi}\sqrt{n+1}\sum_{s=0}^{\frac{n}{2}} \frac{(-1)^s(n-s)!r^{n-2s+1}}{s!\left[(\frac{n}{2}-s)!\right]^2(n-2s+2)}$$

$$F_\theta^i = 0 \tag{1.18}$$

Given the inherent simplicity of modal projection from slope measurements according to eq. 1.7, it is somewhat surprising that these functions have not seen more widespread use

in wavefront sensing and estimation problems. We anticipate that the knowledge of their minimum noise propagation behaviour will make them a simple and attractive approach to wavefront estimation from slope measurements and lead to their increasing use.

It is interesting to note that the problem we have discussed of obtaining a set of auxiliary vector functions which gives minimum wavefront error norm has a close analogy with minimum physical principles. Eq.1.17 may in fact be identically used to describe the equilibrium shape of a membrane (with fixed boundary described by the closed contour C) over which a force $\mathbf{P}_k(\underline{x})$ acts to produce vertical displacement $\lambda(\underline{x})$. In this case, $\int \mathbf{F}_k^2(\underline{x})\, d^2\underline{x}$ represents the potential energy stored in the membrane and will be equal to the work done in bringing the membrane to it's equilibrium position ($\int \mathbf{P}_k(\underline{x})\, \lambda(\underline{x}) d^2\underline{x}$) only in the case of conservative fields - i.e. when $\mathbf{F}_k(\underline{x}) = \nabla\lambda_k(\underline{x})$ [11].

Acknowledgements

The authors would like to acknowledge the support provided for this work by the European Office of Aerospace Research and Development (SPC-98-4050) and the Engineering and Physical Sciences Research Council, U.K. (GR/L26339).

References

[1]. W.H. Southwell, J. Opt. Soc. Am., Vol 70, No. 8, (1980)

[2]. R.H. Hudgin, J. Opt. Soc. Am., Vol 67, 375-378, (1977)

[3]. J. Herrmann, J. Opt. Soc. Am., Vol 70, 28-35, (1980)

[4]. D.L. Fried, J. Opt. Soc. Am.,Vol 67, 370-375, (1977)

[5] E.P. Wallner, J. Opt. Soc. Am., 73 (1983) 1771

[6] C.J. Solomon, J.C. Dainty and N.J. Wooder, Optical Review 2 (1995) 217.

[7] E. Acosta, S. Bará, M.A. Rama and S. Ríos, Opt. Lett. 20 (1995) 1083.

[8] S. Bará, S. Ríos and E. Acosta, J. Opt. Soc. Am A 13 (1996) 1467.

[9] S. Ríos, E. Acosta and S. Bará, Opt. Comm. 133 (1997) 443.

[10] A. Gavrielides, Opt. Lett. 7 (1982) 526.

[11] H.F. Weinberger, "A First Course in Partial Differential Equations", chapter 3, John Wiley and Sons, U.S.A. (1965)

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1999	Final Report

4. TITLE AND SUBTITLE

Simulation Of Low Light Level Adaptive Optics Systems

5. FUNDING NUMBERS

F61775-98-WE079

6. AUTHOR(S)

Dr. Christopher J. Solomon

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Kent
Canterbury CT2 7NR
United Kingdom

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

EOARD
PSC 802 BOX 14
FPO 09499-0200

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

SPC 98-4050

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

A

13. ABSTRACT (Maximum 200 words)

This report results from a contract tasking University of Kent as follows: The contractor will investigate computer modeling of adaptive optics systems operating at low light levels in an attempt to establish their performance limitations.

14. SUBJECT TERMS

EOARD, Modelling & Simulation, Adaptive Optics, Atmospheric Propagation

15. NUMBER OF PAGES

25

16. PRICE CODE

N/A

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL